

Continuity bounds on the quantum relative entropy — II

Koenraad M.R. Audenaert*

*Department of Mathematics, Royal Holloway,
University of London, Egham TW20 0EX, UK*

Jens Eisert

*Dahlem Center for Complex Quantum Systems,
Freie Universität Berlin, D-14195 Berlin, Germany*

(Dated: January 20, 2013)

The quantum relative entropy is frequently used as a distance measure between two quantum states, and inequalities relating it to other distance measures are important mathematical tools in many areas of quantum information theory. We have derived many such inequalities in previous work. The present paper is a follow-up on this, and provides a sharp upper bound on the relative entropy in terms of the trace norm distance and of the smallest eigenvalues of both states concerned. The result obtained here is more general than the corresponding one from our previous work. As a corollary, we obtain a sharp upper bound on the regularised relative entropy introduced by Lendi, Farhadmotamed and van Wonderen.

PACS numbers: 03.65.Hk

I. INTRODUCTION

The quantum relative entropy of states of quantum systems is a measure of how well one quantum state can be operationally distinguished from another and quantifies the extent to which one hypothesis ρ differs from an alternative hypothesis σ in the sense of quantum hypothesis testing [5, 7, 8, 11]. It is defined as

$$S(\rho||\sigma) = \text{Tr } \rho(\log \rho - \log \sigma)$$

*Electronic address: koenraad.audenaert@rhul.ac.uk

for states ρ and σ [10] whenever the support of ρ is contained in the support of σ , and is defined to be $+\infty$ otherwise.

In [3] we presented a number of inequalities relating the quantum relative entropy, used as a distance measure, to the trace norm distance. The present paper is a follow-up on this work, and concerns a sharp upper bound on the relative entropy $S(\rho||\sigma)$ in terms of the trace norm distance $||\rho - \sigma||_1/2$, when the smallest eigenvalues of ρ and σ are given. The need for these smallest eigenvalues stems from the fact that the relative entropy can be infinite when the kernel of σ is not contained in the kernel of ρ . Rastegin obtained similar inequalities for the relative q -entropy [9].

As a special case of the main theorem proven here (Theorem 1), we reobtain Theorem 6 of [3]. The proof given in [3] was incorrect, and the proof we give here serves as a correction and at the same time as a generalisation.

We also obtain an upper bound (Corollary 2) on the so-called regularised relative entropy, introduced by Lendi et. al. [6] as one possible means to circumvent the problem of infinities of the ordinary relative entropy. The regularised relative entropy is defined as

$$R(\rho||\sigma) = c_d S(\rho + \mathbf{I}||\sigma + \mathbf{I}),$$

where c_d is a certain normalisation constant depending on d , the dimension of state space. Note that $S(\rho + \mathbf{I}||\sigma + \mathbf{I}) \leq \log 2$, with equality for orthogonal pure states, hence one could also choose the normalisation constant to be $1/\log 2$.

In the following section, we introduce the notations and mathematical tools necessary for the proofs. Then, in Section III, we derive an integral representation for the relative entropy between non-normalised states (i.e. positive definite matrices), which is also essential for the proofs. An upper bound on the relative entropy for non-normalised states is derived in Section IV, which is then used in Section V to obtain the promised sharp upper bound on the relative entropy for normalised states.

II. NOTATIONS AND PRELIMINARIES

In this paper we will work exclusively in finite dimensional Hilbert spaces, so that quantum states can be represented by positive semidefinite matrices. We denote the identity matrix by \mathbf{I} , and scalar matrices $a\mathbf{I}$ simply by a (for $a \in \mathbf{R}$) when no confusion can arise. The matrix norms $||\cdot||_1$ and $||\cdot||_\infty$ are the trace norm and operator norm, respectively.

The von Neumann entropy can be defined for positive definite matrices as

$$S(A) = -\text{Tr } A \log A, \quad (1)$$

which coincides with the usual definition for density matrices. Furthermore, we define $S(0) = 0$.

Likewise, the quantum relative entropy can be defined for positive definite matrices A and B as

$$S(A||B) = \text{Tr } A(\log A - \log B). \quad (2)$$

This definition still holds for positive semidefinite A and B provided the support of B is contained in the support of A ; otherwise one defines $S(A||B) = +\infty$. The quantum relative entropy satisfies the scaling property

$$S(aA||aB) = S(A||B), a > 0. \quad (3)$$

The logarithm appearing in (1) and (2) is the matrix logarithm. For $x > 0$, we have the following integral representation for the scalar logarithm:

$$\log x = \int_0^\infty ds \left(\frac{1}{1+s} - \frac{1}{x+s} \right). \quad (4)$$

and for $A > 0$ we define the matrix logarithm as

$$\log A = \int_0^\infty ds \left(\frac{1}{1+s} - (A+s)^{-1} \right). \quad (5)$$

The methods we will use require the derivative of the matrix logarithm; see also [1, 2]. From the integral representation of the logarithm we get, for $A > 0$,

$$\left. \frac{d}{dt} \right|_{t=0} \log(A + t\Delta) = \int_0^\infty ds (A+s)^{-1} \Delta (A+s)^{-1}.$$

As is customary, we define the following linear map for $A > 0$:

$$\mathcal{T}_A(\Delta) = \int_0^\infty ds (A+s)^{-1} \Delta (A+s)^{-1}. \quad (6)$$

Thus

$$\left. \frac{d}{dt} \right|_{t=0} \log(A + t\Delta) = \mathcal{T}_A(\Delta). \quad (7)$$

Again, (6) and (7) are also valid for $A \geq 0$ provided $\ker A \subseteq \ker \Delta$.

From the integral representation of \mathcal{T} it follows that, for any $A > 0$, \mathcal{T}_A preserves the positive semidefinite order: if $X \leq Y$, then $\mathcal{T}_A(X) \leq \mathcal{T}_A(Y)$.

For $x > 0$, the integral $\int_0^\infty ds \, x/(x+s)^2$ is equal to 1. Hence, for $A > 0$,

$$\mathcal{T}_A(A) = \int_0^\infty ds \, (A+s)^{-1} A (A+s)^{-1} = \mathbf{I}. \quad (8)$$

An argument that we will use frequently is the special structure of the Jordan decomposition of a traceless Hermitian matrix. Let Δ be Hermitian, and $\text{Tr } \Delta = 0$. The Jordan decomposition of Δ is given by $\Delta = \Delta_+ - \Delta_-$, where Δ_+ and Δ_- are positive semidefinite and mutually orthogonal, i.e. $\Delta_+ \Delta_- = 0$. We have $\text{Tr } \Delta = \text{Tr } \Delta_+ - \text{Tr } \Delta_-$, hence $\text{Tr } \Delta_+ = \text{Tr } \Delta_-$. Thus

$$\|\Delta\|_1 = \text{Tr } \Delta_+ + \text{Tr } \Delta_- = 2 \text{Tr } \Delta_+. \quad (9)$$

It will also be clear that $\|\Delta\|_\infty$ is bounded above by $\text{Tr } \Delta_+$, and thus

$$\|\Delta\|_\infty \leq \|\Delta\|_1/2, \quad (10)$$

whenever Δ is traceless and Hermitian.

III. AN INTEGRAL REPRESENTATION OF THE RELATIVE ENTROPY

In this section we derive an integral representation of the quantum relative entropy for non-normalised states,

$$S(A||B) = \text{Tr } A(\log A - \log B),$$

in terms of a differentiable path $s \mapsto C(s)$, where $C(0) = A$, $C(1) = B$ and $C(s) > 0$.

Lemma 1 *Let $A \geq 0$ and $B > 0$. Let $0 \leq s \leq 1$ and let $s \mapsto C(s)$ be a continuous, differentiable path joining A and B (that is, $C(0) = A$ and $C(1) = B$) and for all $s \in (0, 1)$, $C(s) > 0$. Then the relative entropy $S(A||B)$ has the following integral representation:*

$$S(A||B) = \text{Tr}(A - B) + \int_0^1 ds \, \text{Tr} \frac{dC}{ds} (\log B - \log C(s)). \quad (11)$$

In particular, if the path is linear, $C(s) = (1-s)A + sB$, then (11) becomes

$$S(A||B) = \text{Tr}(A - B) + \text{Tr}(B - A) \log B - \int_0^1 ds \, \text{Tr}(B - A) \log(A + s(B - A)). \quad (12)$$

Proof. First we rewrite $S(A||B)$ as

$$\begin{aligned} S(A||B) &= \text{Tr } A(\log A - \log B) \\ &= \text{Tr}(B - A) \log B + (S(B) - S(A)). \end{aligned}$$

Both terms can be written as integrals. For the first term we have

$$\text{Tr}(B - A) \log B = \text{Tr} \int_0^1 ds \frac{dC}{ds} \log B.$$

The second term can be written as:

$$S(B) - S(A) = -\text{Tr}(B - A) - \int_0^1 ds \text{Tr} \frac{dC}{ds} \log C(s).$$

This can be shown as follows. Let $f(x) = -x \log x$ be the function defining the von Neumann entropy $S(A) = \text{Tr} f(A)$. As $f'(x) = -1 - \log x$, we have, for $0 < s < 1$,

$$\begin{aligned} \frac{d}{ds} S(C(s)) &= \frac{d}{ds} \text{Tr} f(C(s)) \\ &= \text{Tr} f'(C(s)) \frac{dC}{ds} \\ &= -\text{Tr} \frac{dC}{ds} - \text{Tr} \frac{dC}{ds} \log C(s). \end{aligned}$$

Integrating over s in the interval $[0, 1]$ yields, indeed,

$$\begin{aligned} S(B) - S(A) &= S(C(1)) - S(C(0)) \\ &= \int_0^1 ds \frac{d}{ds} S(C(s)) \\ &= -\int_0^1 ds \text{Tr} \frac{dC}{ds} - \int_0^1 ds \text{Tr} \frac{dC}{ds} \log C(s) \\ &= -\text{Tr}(B - A) - \int_0^1 ds \text{Tr} \frac{dC}{ds} \log C(s). \end{aligned}$$

Finally, adding the two integral representations yields (11). □

IV. UPPER BOUND ON THE RELATIVE ENTROPY FOR NON-NORMALISED STATES

In this section, we prove the main technical proposition (Proposition 1) from which the promised sharp bound will follow. The proposition provides an upper bound on the relative entropy for non-normalised states A and B with equal trace, in terms of the trace norm distance T between A and B , and of the minimal eigenvalues of A and B .

We will denote the lowest eigenvalue of A by α , and the lowest eigenvalue of B by β . First we establish the allowed range of T in terms of α and β . It turns out that the trace norm distance between A and B cannot be smaller than $|\alpha - \beta|$:

Lemma 2 *Let A, B be positive semidefinite $n \times n$ matrices with $\text{Tr } A = \text{Tr } B$, and $\lambda_{\min}(A) = \alpha$ and $\lambda_{\min}(B) = \beta$. Then $T = \|A - B\|_1/2 \geq |\alpha - \beta|$.*

Proof. We assume first that $\alpha \geq \beta$. Let $\Delta := A - B$ have Jordan decomposition $\Delta = \Delta_+ - \Delta_-$. Since $\text{Tr } A = \text{Tr } B$, we have $\text{Tr } \Delta = 0$, hence $\|A - B\|_1 = 2 \text{Tr } \Delta_+$.

Denoting the vector of eigenvalues sorted in non-increasing order by the symbol λ^\downarrow , we then clearly have

$$\|A - B\|_1 = 2 \text{Tr } \Delta_+ \geq 2\lambda_1^\downarrow(\Delta_+) = 2\lambda_1^\downarrow(A - B).$$

Now, by Lidskii's Theorem (e.g. inequality (III.12) in [4]), for all Hermitian A and B , the vector $\lambda^\downarrow(A) - \lambda^\downarrow(B)$ is majorised by the vector $\lambda^\downarrow(A - B)$. In particular,

$$\lambda_1^\downarrow(A - B) \geq \max_j \{\lambda_j^\downarrow(A) - \lambda_j^\downarrow(B)\} \geq \lambda_n^\downarrow(A) - \lambda_n^\downarrow(B).$$

By the hypothesis of the lemma, the last expression is equal to $\alpha - \beta$.

Hence we have shown that $\|A - B\|_1 \geq 2(\alpha - \beta)$ when $\alpha - \beta \geq 0$. When $\alpha - \beta \leq 0$ we can just swap the roles of A and B and obtain $\|A - B\|_1 \geq 2(\beta - \alpha)$. \square

Because of the scaling property (3) we can restrict ourselves to the case $\beta = 1$.

Proposition 1 *Let A, B be positive definite with $\text{Tr } A = \text{Tr } B$, $\lambda_{\min}(A) = \alpha$, $\lambda_{\min}(B) = 1$ and $T := \|A - B\|_1/2$. Then $T \geq |\alpha - 1|$, and Then, for $\alpha > 0$,*

$$S(A||B) \leq (1 + T) \log(1 + T) - \alpha \log(1 + T/\alpha), \quad (13)$$

where $\alpha \mapsto -\alpha \log(1 + T/\alpha)$ is monotone decreasing, and $-\alpha \log(1 + T/\alpha) =: 0$ for $\alpha = 0$.

Moreover, quality can be achieved for any allowed values of α and T .

The proof relies on the following lemma:

Lemma 3 *Let a and b be two positive definite matrices with $\text{Tr } a = \text{Tr } b$, and let $t = \|b - a\|_1/2$. If $a \geq \gamma$, with γ a non-negative scalar, then*

$$\mathcal{T}_b(b - a) \leq \frac{t}{\gamma + t}. \quad (14)$$

Proof. Let $\delta = b - a$, which is Hermitian with trace equal to 0 and trace norm equal to $2t$. Thus, by (10), $\|\delta\|_\infty \leq t$, or $\delta \leq t$. We also have $t\gamma \leq ta$. Combining the two inequalities yields $\gamma\delta \leq ta = t(b - \delta)$. Hence,

$$b - a = \delta \leq \frac{t}{\gamma + t}b.$$

Since the operator $Y \mapsto \mathcal{T}_X(Y)$ is order-preserving for $X > 0$, applying this operator to both sides yields

$$\mathcal{T}_b(b - a) \leq \frac{t}{\gamma + t}\mathcal{T}_b(b) = \frac{t}{\gamma + t}.$$

□

Proof of Proposition 1. We consider strictly positive α first. Let us apply Lemma 3 to the case $a = A$ and $b = A + x(B - A)$, with A and B the matrices of the proposition and $0 \leq x \leq 1$. Let $\Delta = B - A$. Then $\delta = x\Delta$, $t = xT$ and $\gamma = \alpha$. By the lemma, we then have (after dividing by x)

$$\mathcal{T}_{A+x\Delta}(\Delta) \leq \frac{T}{\alpha + xT}. \quad (15)$$

Likewise, by setting $a = B$, $b = B + (1 - x)(A - B)$ and $\gamma = 1$, we get

$$\mathcal{T}_{B+(1-x)(-\Delta)}(-\Delta) \leq \frac{T}{1 + (1 - x)T}.$$

Noting that $B - (1 - x)\Delta = xB + (1 - x)A = A + x\Delta$, this yields the lower bound

$$\mathcal{T}_{A+x\Delta}(\Delta) \geq \frac{-T}{1 + (1 - x)T}. \quad (16)$$

Again we exploit the Jordan decomposition of Δ , $\Delta = \Delta_+ - \Delta_-$ with $\Delta_+, \Delta_- \geq 0$ and $\text{Tr } \Delta_+ = \text{Tr } \Delta_- = T$, due to the facts that $\text{Tr } \Delta = 0$ and $\|\Delta\|_1 = 2T$. Combining this with (15) and (16), and exploiting the fact that for $X \geq 0$, $Y \leq y$ implies $\text{Tr } XY \leq y \text{Tr } X$, we get

$$\begin{aligned} \text{Tr } \Delta \mathcal{T}_{A+x\Delta}(\Delta) &= \text{Tr } \Delta_+ \mathcal{T}_{A+x\Delta}(\Delta) - \text{Tr } \Delta_- \mathcal{T}_{A+x\Delta}(\Delta) \\ &\leq T \left(\frac{T}{\alpha + xT} - \frac{-T}{1 + (1 - x)T} \right). \end{aligned}$$

Now let s be a scalar, $0 \leq s \leq 1$. Integrating the previous inequality over x from s to 1 yields

$$\text{Tr } \Delta (\log(A + \Delta) - \log(A + s\Delta)) \leq T(\log(\alpha + T) - \log(\alpha + sT) + \log(1 + (1 - s)T)).$$

Integrating a second time, now over s from 0 to 1, yields:

$$\int_0^1 ds \text{Tr } \Delta (\log B - \log(A + s\Delta)) \leq (1 + T) \log(1 + T) + \alpha(\log \alpha - \log(\alpha + T)).$$

To finish the proof, we define the rectilinear path $C(s) = sB + (1 - s)A$, for which $dC/ds = B - A = \Delta$, and use Lemma 1 to show that the left-hand side is just $S(A||B)$.

The strict positivity of α is required to satisfy the conditions of Lemma 3. However, by continuity of the relative entropy in its first argument, the bound must be valid for $\alpha = 0$ too. In the limit of α tending to 0, $\alpha(\log \alpha - \log(\alpha + T))$ goes to 0.

Finally, we show that equality can be obtained for every allowed value of T and α . Indeed, taking

$$A = \begin{pmatrix} 1+T & 0 \\ 0 & \alpha \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 0 \\ 0 & T+\alpha \end{pmatrix}$$

satisfies all the requirements of the proposition and yields equality in (13). \square

V. SHARP UPPER BOUNDS ON THE RELATIVE ENTROPY AND REGULARISED RELATIVE ENTROPY

Proposition 1 allows us to derive an upper bound on the ordinary relative entropy between density operators ρ and σ when the eigenvalues of ρ and σ are bounded below by the values α and β , respectively.

Theorem 1 *Consider density matrices ρ and σ , with smallest eigenvalues $\lambda_{\min}(\rho) = \alpha$ and $\lambda_{\min}(\sigma) = \beta$. Then $T := \|\rho - \sigma\|_1/2 \geq |\alpha - \beta|$ and, for $\alpha, \beta > 0$,*

$$S(\rho||\sigma) \leq (\beta + T) \log(1 + T/\beta) - \alpha \log(1 + T/\alpha), \quad (17)$$

and, in the limit $\alpha \rightarrow 0$,

$$S(\rho||\sigma) \leq (\beta + T) \log(1 + T/\beta). \quad (18)$$

Proof. We use the scaling property and Proposition 1, with $A = \rho/\beta$ and $B = \sigma/\beta$. The formula of Proposition 1 can be taken over completely by replacing α by α/β , T by T/β , and multiplying the right-hand side of each bound by β . \square

Note that, because of the extra normalisation requirement $\text{Tr } \rho = \text{Tr } \sigma = 1$, equality can now only be achieved for states of dimension at least 3.

If α is not specified, we must take the maximum of $(\beta + T) \log(1 + T/\beta) - \alpha \log(1 + T/\alpha)$ over all allowed values of α , with β and T kept fixed. In doing so we retrieve Theorem 6 of [3]. The proof given here supplants the incorrect proof in the published version of [3].

Corollary 1 Consider density matrices ρ and σ , where σ has smallest eigenvalue $\lambda_{\min}(\sigma) = \beta$. Let $T := \|\rho - \sigma\|_1/2$. If $T \leq \beta$

$$S(\rho||\sigma) \leq (\beta + T) \log(1 + T/\beta) + (\beta - T) \log(1 - T/\beta), \quad (19)$$

and if $T \geq \beta$,

$$S(\rho||\sigma) \leq (\beta + T) \log(1 + T/\beta). \quad (20)$$

Proof. Let $\lambda_{\min}(\rho) = \alpha$. To find an upper bound on $S(A||B)$ in the case that α is not specified, we maximise the bound (17) over all allowed α . Because of Lemma 2, $T \geq |\alpha - \beta|$. Hence, the range of α is $[\max(0, \beta - T), \beta + T]$. The quantity to be maximised is $-\alpha \log(1 + T/\alpha)$, which is monotonously decreasing in α . Thus, its maximum is attained for the minimally allowed α , being $\max(0, \beta - T)$. The two cases of the corollary follow. \square

We immediately obtain an upper bound on the regularised relative entropy in terms of the trace norm distance.

Corollary 2 For d -dimensional density matrices ρ and σ , with smallest eigenvalues $\lambda_{\min}(\rho) = \alpha$ and $\lambda_{\min}(\sigma) = \beta$ and $T := \|\rho - \sigma\|_1/2$,

$$R(\rho||\sigma) := c_d S(\rho + \mathbf{I}||\sigma + \mathbf{I}) \quad (21)$$

$$\leq c_d ((\beta + 1 + T) \log(1 + T/(\beta + 1)) - (\alpha + 1) \log(1 + T/(\alpha + 1))) \quad (22)$$

$$\leq c_d T \log(1 + T). \quad (23)$$

From dimension 3 onwards, inequality (23) is sharp. Equality can be achieved for any allowed value of T , by the diagonal states $\rho = \text{Diag}(1 - t, t, 0)$ and $\sigma = \text{Diag}(1 - t, 0, t)$, where t can be any number between 0 and 1.

Proof. Upper bound (22) is a straightforward application of Theorem 1 (apart from a rescaling of $\rho + \mathbf{I}$ and $\sigma + \mathbf{I}$, which has no effect on the bound itself). When no information about α and β is known one can use the bound (23) which follows by exploiting the fact that both $(\beta + 1 + T) \log(1 + T/(\beta + 1))$ and $-(\alpha + 1) \log(1 + T/(\alpha + 1))$ are monotonically decreasing, hence expression (22) is maximal for $\alpha = \beta = 0$. \square

Acknowledgments

This work was supported by the European Commission (Qessence, Compas, Minos) and the European Research Councils (EURYI). We gratefully acknowledge the referee for many invaluable comments.

-
- [1] K.M.R. Audenaert, “Telescopic Relative Entropy”, arXiv:1102.3040. Presented at TQC2011, May 22 2011, Madrid (2011).
 - [2] K.M.R. Audenaert, “Telescopic Relative Entropy–II: Triangle Inequalities”, arXiv:1102.3041 (2011).
 - [3] K.M.R. Audenaert and J. Eisert, J. Math. Phys. **46**, 102104 (2005).
 - [4] R. Bhatia, *Matrix Analysis*, Springer, Heidelberg (1997).
 - [5] F. Hiai and D. Petz, “The proper formula for relative entropy and its asymptotics in quantum probability”, Comm. Math. Phys. **143**, 99–114 (1991).
 - [6] K. Lendi, F. Farhadmotamed and A.J. van Wonderen, J. Stat. Phys. **92**(5/6), 1115–1135 (1998).
 - [7] T. Ogawa and H. Nagaoka, “Strong converse and Stein’s lemma in quantum hypothesis testing”, IEEE Trans. Inform. Theory **47**, 2428–2433 (2000).
 - [8] M. Ohya and D. Petz, *Quantum Entropy and its Use*, Springer, Heidelberg (1993).
 - [9] A.E. Rastegin, “Upper continuity bounds on the relative q -entropy for $q > 1$ ”, arXiv:1010.1355 (2010).
 - [10] H. Umegaki, Kodai Math. Sem. Rep. **14**, 59 (1962).
 - [11] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).